

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

## La standardizzazione italiana del test ToM Storybooks: dati preliminari

### This is the author's manuscript

*Original Citation:*

*Availability:*

This version is available <http://hdl.handle.net/2318/126605> since

*Published version:*

DOI:10.1421/73146

*Terms of use:*

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)



# UNIVERSITÀ DEGLI STUDI DI TORINO

***This is an author version of the contribution published on:***

*Questa è la versione dell'autore dell'opera:*

**Molina, P., Bulgarelli, D. (2012), La standardizzazione italiana del test ToM Storybooks: dati preliminari, *Giornale Italiano di Psicologia*, XXXIX (3), 747-762 [DOI: 10.1421/73146]**

***The definitive version is available at:***

*La versione definitiva è disponibile alla URL:*

<https://www.mulino.it/riviste/issn/0390-5349>

## **La standardizzazione italiana del test *ToM Storybooks*: dati preliminari**

### Riassunto

Il *ToM Storybooks* (Blijd-Hoogewys, van Geert, Serra e Minderaa, 2008) è un test olandese che valuta le molteplici competenze coinvolte nella comprensione della mente (Teoria della Mente), attraverso misure diverse per ogni componente. Il campione italiano di standardizzazione consiste di 446 bambini (223 femmine) a sviluppo tipico di età compresa tra 3 e 8 anni. L'andamento delle risposte cresce regolarmente con l'età. Lo strumento si è rivelato affidabile. Non sono emerse differenze di genere, né alla presenza di fratelli, o al fatto di avere genitori stranieri o di parlare più lingue in famiglia.

### Abstract

The Italian standardization of the *ToM Storybooks* test

The *ToM Storybooks* (Blijd-Hoogewys, van Geert, Serra & Minderaa, 2008) is a Dutch test evaluating different competences involved in mental state understanding (Theory of Mind), using different measures for each component. The Italian standardization sample consists of 466 typically developing children (223 females; age: 3-8 years). The response trend grows regularly with age. The tool proved to be reliable. No differences emerged in respect of gender, presence of siblings, having a foreign parent or talking different languages at home.

### Parole chiave:

Theory of Mind, test, children, standardization, *ToM Storybooks*

## 1. Introduzione

La Teoria della Mente (ToM) è la capacità di attribuire stati mentali a sé se stessi e agli altri, e di comprendere che il comportamento delle persone è prevedibile e spiegabile attraverso gli stati interni (Mitchell, 1997). Nella visione costruttivista di Wellman (1990), i bambini in età prescolare possiedono già una complessa teoria “naïf” per spiegare ciò che regola le azioni umane, mettendo in gioco diversi fenomeni. In questo schema di comprensione della mente, le azioni umane sono guidate da credenze e desideri, che sono fenomeni soggettivi e interni; essi sono ricorsivi, nel senso che una credenza ne genera un'altra e un desiderio ne genera un altro a sua volta. Le credenze sono influenzate dalle percezioni (visive, uditive, olfattive, tattili, ecc.), che ci permettono di conoscere il mondo esterno in un certo modo. I desideri, in questo schema semplificato, sono invece influenzati dagli stati emotivi e da quelli fisiologici (e.g., fame, sete, dolore, eccitazione). Le azioni sono determinate dagli stati interni (credenze e desideri), e comportano a loro volta delle reazioni, generalmente di tipo emotivo (e.g., la soddisfazione di un desiderio genera piacere o contentezza, così come qualcosa di atteso può non avverarsi e quindi generare sorpresa).

Nell'impostazione teorica di Wellman, questo schema semplificato descrive la Teoria della Mente che possiedono i bambini in età prescolare, a partire dai 3 anni di età circa. La ToM è quindi una competenza molto complessa e sfaccettata, che viene concettualizzata come composta da sotto-competenze diverse. Tra di esse, la capacità di utilizzare e comprendere correttamente le credenze e i desideri richiede che i bambini si rendano conto che questi fenomeni mentali sono caratterizzati dalla proprietà logica dell'opacità referenziale (Battistelli, 1997). Essa consiste nel fatto che credenze e desideri non sono “descrizioni” oggettive di una realtà esterna e quindi necessariamente descrizioni vere o false, ma sono piuttosto “descrizioni” relative a uno stato interno: indipendentemente dalla realtà oggettiva a cui si riferisce, la credenza può essere vera o falsa a seconda che

l'individuo creda o non creda. Analogamente, ad esempio, l'ottenere una palla in regalo non rende automaticamente felice una persona, perché questo dipende dal fatto che essa desideri o meno la palla. Questo aspetto genera, nei bambini, problemi di inferenza sulla veridicità delle credenze e dei desideri, che vengono generalmente superati intorno ai 4 anni di età. Classicamente, nei bambini in età prescolare la ToM è stata valutata principalmente proprio attraverso la comprensione della Falsa credenza (FC – Wimmer e Perner, 1983; Baron-Cohen, Leslie e Frith, 1985). Le prove messe a punto sono pensate per osservare non l'insieme delle componenti che costituiscono la ToM, bensì questo specifico aspetto, che richiede al bambino di avere chiare le implicazioni dell'opacità referenziale. La comprensione della FC si è rivelata una competenza generalizzabile, poiché l'andamento delle risposte dei bambini non cambia, anche in contesti culturali estremamente diversi; inoltre l'andamento evolutivo rimane simile nelle diverse ricerche, anche se le prove utilizzate per valutare la FC sono molto differenti tra loro, ad esempio rispetto ai tipi di compito proposto, al tipo di domande, alla natura del protagonista o dell'oggetto nascosto, alla sua reale presenza, alla salienza dello stato mentale del protagonista, ecc. (Wellman, Cross e Watson, 2001). In letteratura, inoltre, sono stati indagati diversi aspetti che possono influenzare la prestazione dei bambini alle prove di FC. La relazione tra linguaggio e ToM è ormai associata (Milligan, Astington e Dack, 2007) e i bambini bilingui sembrano presentare prestazioni migliori (Kovacs, 2009). Non vi è accordo, invece, circa la presenza di differenze di genere: alcuni studi non riportano differenze significative (cf. Renouf et al., 2010) e, quando esse vengono rilevate, sono di solito a favore delle bambine (Charman et al., 2002; Walker, 2005). Anche la presenza di fratelli sembra legata a migliori prestazioni nelle prove di FC (McAlister e Peterson, 2007): sarebbe la possibilità di interagire con bambini più grandi a stimolare una maggiore competenza socio-cognitiva, tant'è che anche i bambini che nella scuola dell'infanzia frequentano classi verticali, dove sono compresenti bambini di

diverse età, sono avvantaggiati nei compiti di ToM rispetto ai bambini che frequentano classi orizzontali, che sono composte da bambini della stessa età (Wang e Su, 2009).

Tuttavia, anche rispetto alla presenza di fratelli non vi è un univoco accordo e alcuni studi non rilevano differenze significative (Das e Babu, 2004).

Come si è visto, la letteratura presente è sostanzialmente concentrata sullo studio della comprensione della FC, ma essa è solo una delle componenti della ToM. È utile poterne ottenere, invece, una misura complessa, che tenga conto delle diverse competenze che entrano in gioco nella comprensione della mente. Una misura di questo tipo risulta, in quanto composita, maggiormente affidabile (Hughes, Adlam, Happé, Jackson, Taylor e Caspi, 2000), e questo presenta dei vantaggi sia a livello teorico, come messo in luce dalla Teoria dei Sistemi Dinamici (van Geert, 2003) sia a livello clinico, dove la valutazione sul singolo deve essere il più possibile attendibile.

Pochi strumenti permettono misure articolate, ottenute da prove diverse che valutino le varie componenti che entrano in gioco nella comprensione della mente e non solo la FC: la ToM Battery (Happé, 1994), il ToM-Test (Steerneman, Meesters e Muris, 2002), i ToM Tasks di Tager-Flusberg (2003) e i ToM Tasks di Wellman e Liu (2004) sono tra questi. A differenza del test di Wellman e Liu, che valuta aspetti semplici della ToM, gli altri tre test sono pensati anche per i bambini in età scolare e presentano prove che valutano aspetti evoluti della ToM.

La ToM Battery (Happé, 1994) è utilizzabile a partire dai 6 anni ai 10 anni e include compiti più semplici (finzione, distinzione apparenza-realtà) e altri più evoluti (FC di secondo ordine, sarcasmo, menzogna) della ToM. Le prove di FC di primo e secondo ordine sono quelle classiche, mentre la parte originale del test consiste in 24 Strange Stories, costituite da vignette illustrate e relativi racconti.

Il ToM-Test (Muris, Steerneman, Meesters, Merckelbach, Horselenberg, van den Hogen, e van Dongen, 1999; Steerneman, et al., 2002) è basato sull'impostazione teorica di Flavell, Miller e Miller (1993). Pensato per i bambini tra i 3 e i 12 anni, esso presenta prove ordinate per difficoltà di superamento, relative a precursori della ToM (riconoscimento delle emozioni, comprensione del fare finta), ad aspetti specifici della ToM (la credenza di primo ordine – “io penso che” – e di falsa credenza), e ad aspetti più complessi (come la credenza di secondo ordine – “io penso che tu pensi” – e di comprensione dello humor). Le prove sono proposte sotto forma di storie illustrate.

I ToM Tasks di Tager-Flusberg (2003) sono stati pensati per osservare la ToM in bambini autistici di età compresa tra i 4 e i 14 anni e coprono anche competenze che, nello sviluppo tipico, emergono in tarda età scolare. I compiti riguardano competenze precoci di ToM (fare finta, desiderio), medie (ruolo del percezione sulla conoscenza, falsa credenza e sticker hiding) e avanzate (credenza di secondo ordine, bugie e scherzi, tratti di personalità e l'impegno morale).

I ToM Tasks di Wellman e Liu (2004) sono pensati per bambini dai 3 ai 6 anni. La scelta delle sette prove che ne fanno parte è stata basata su una meta-analisi degli studi presenti in letteratura, avente la finalità di identificare quali tipi di prove meglio permettessero di differenziare la competenza di ToM in età prescolare. Le prove sono presentate sotto forma di storie con l'ausilio di pupazzetti; utilizzando per l'analisi le scale di Rash, esse sono risultate ordinate per difficoltà crescente: desiderio diverso, credenza diversa, accesso alla conoscenza, falsa credenza, falsa credenza esplicita, ruolo della credenza sull'emozione e nascondere l'emozione; per una descrizione delle prove, si veda l'appendice in Wellman e Liu (2004).

In questo quadro, si colloca anche il *ToM Storybooks* (Blijd-Hoogewys, van Geert, Serra e Minderaa, 2008; Molina e Bulgarelli, 2008; Molina, Bulgarelli e Arati, 2008).

Questo test è stato sviluppato originariamente in lingua olandese per valutare la ToM nei bambini a sviluppo tipico tra i 3 e gli 8 anni; attualmente, oltre alla versione italiana qui presentata, ne esistono anche una versione francese (Thommen, Cartier-Nelles, Wiesendanger, Guidoux, e Blijd-Hoogewys, 2010) e una finlandese (Vesterinen, 2008). Il *ToM Storybooks* presenta prove relative allo sviluppo delle competenze nei bambini in età prescolare, e valuta in modo attento le forme iniziali della costruzione della ToM. Inoltre si propone di fornire un punteggio standardizzato delle competenze relative alla comprensione degli stati mentali, rendendo più affidabile la valutazione sia a livello clinico che di ricerca.

L'obiettivo del nostro lavoro è quindi quello di presentare i dati della standardizzazione italiana di questo strumento.

## 2. Metodo

### 2.1 Descrizione dello strumento

Il *ToM Storybooks* è un test comprensivo, concepito in accordo con la teoria di Wellman (1990): questo si riflette nella scelta delle prove che sono state inserite nel test. Esse, infatti, sono pensate per valutare le componenti della ToM che abbiamo discusso nell'introduzione:

1) Il riconoscimento delle emozioni di base (felicità, tristezza, rabbia, paura e sorpresa; per un esempio, si veda Tabella 1).

2) La comprensione della distinzione mentale/fisico, in specifico:

- La comprensione della specificità dei fenomeni mentali, che sono intangibili, interni, privati, e soggettivi; in particolare, vengono indagati il far finta, il sognare, il pensare e il ricordare (per un esempio, si veda Tabella 1).
- La comprensione del funzionamento di particolari fenomeni fisici, come il fumo o l'odore, che condividono alcune qualità degli stati mentali, come il fatto di non poter essere toccate, percepite con tutti i sensi, ecc. (in inglese *close impostors*).



- La comprensione della distinzione immaginario/reale (posso pensare o sognare una cosa che non esiste?).

3) La comprensione che è necessario percepire per conoscere

4) La comprensione del ruolo dei desideri e del ruolo delle credenze, che

- predicono emozioni;
- predicono comportamenti (per un esempio, si veda Tabella 1).

Tabella 1

Esempi di prove dal ToM Storybooks

Competenza valutata	Esempio di prova
<i>Riconoscimento delle emozioni</i>	Simone ha vinto giocando a biglie. Ha vinto la biglia più bella. Scegli la faccina che va bene. Come ti sembra Simone? Perché Simone ti sembra [emozione]?
<i>Distinzione mentale/fisico: specificità dei fenomeni mentali</i>	Simone guarda le galline. Ce ne sono un mucchio. C'è anche il suo papà. Le galline hanno deposto le uova. Sono nel cesto [apri la finestrella]. [indicare] Intanto la zia Luisa é in cucina [indicare] e anche lo zio Bruno. All'improvviso la zia Luisa ricorda che ha lasciato le uova nel pollaio. Pensa: "Mamma mia, ho dimenticato le uova". Chi non può toccare le uova con le mani adesso? Simone o la zia Luisa? Chi può prendere il cesto con le uova adesso? La zia Luisa o Simone? Perché ... [Simone/la zia Luisa] può prendere le uova? Simone ha il cesto con le uova. Anche il papà può davvero toccare le uova? Chi può adesso nascondere le uova e riprenderle più tardi? Simone o la zia Luisa?
<i>Ruolo della credenza sul comportamento</i>	Simone va a dar da mangiare alle mucche insieme allo zio Bruno. "Puoi prendere il secchio con il mangime?", chiede lo zio Bruno. Il secchio con il mangime può essere [indicare] dietro l'albero o [indicare] nella casetta. Simone vuole trovare il secchio. Pensa che il secchio non sia dietro l'albero. Dove guarderà Simone? Perché Simone guarda ... [là]? Dove Simone pensa che sia il secchio?

Un secondo importante riferimento teorico per il *ToM Storybooks* è quello alla relativo alla Teoria dei Sistemi Dinamici (Ford e Lerner, 1995; Thelen e Smith, 1994; Fischer e Paré-Blagoev, 2000; van Geert, 2003). Questa prospettiva si propone specificamente di analizzare gli elementi di continuità e discontinuità nello sviluppo, ed è quindi particolarmente attenta a quei processi che presentano aspetti di cambiamento qualitativo, così come avviene per la comprensione degli stati mentali: in particolare, la comparsa della comprensione della falsa credenza è un tema canonico del dibattito intorno alla continuità/discontinuità dei processi mentali.

Rispetto agli altri quattro test comprensivi che abbiamo precedentemente descritto (Happé, 1994; Steerneman, et al., 2002; Tager-Flusberg, 2003; Wellman e Liu, 2004), il *ToM Storybooks* è specificamente pensato per la valutazione di aspetti della ToM che emergono in età prescolare (e si raffinano ancora in età scolare): non presenta compiti complessi come la credenza di second'ordine o la comprensione dello humor, ma approfondisce aspetti relativi, ad esempio, alla distinzione mentale-fisico, reale-immaginario e i *close impostors*; queste competenze relative alla distinzione di cosa sia un fenomeno mentale rispetto a un fenomeno fisico sono alla base della comprensione del funzionamento specifico degli stati mentali e non sono indagati dagli altri test comprensivi.

L'aspetto peculiare dei test comprensivi risiede nel fatto che la valutazione della competenza del bambino è più accurata, poiché egli ha un'ampia possibilità di espressione e la sua prestazione risulta meno influenzata da aspetti contingenti delle singole prove. Nel *ToM Storybooks*, oltre a questo, alcune prove create per valutare una specifica competenza sono presentate più volte (al massimo quattro) ma in contesti di storie differenti: questo permette di differenziare meglio la prestazione dei bambini che hanno una competenza ormai stabilizzata rispetto a coloro che non ce l'hanno ancora e, come sottolineato dalla

Teoria dei Sistemi Dinamici, questa modalità di valutazione è particolarmente utile per incrementare la validità. In totale il *ToM Storybooks* è costituito quindi da 34 prove, che valutano le componenti principali della ToM nel contesto di storie e situazioni quotidiane.

Come emerge dagli esempi riportati in Tabella 1, il *ToM Storybooks* comprende due tipi di domande: domande sì/no e domande aperte. Il test prevede quindi il calcolo di un triplice punteggio:

- punteggio quantitativo (max 77 punti): relativo alle 77 risposte chiuse;
- punteggio qualitativo (max 36 punti): relativo alle 18 risposte aperte, ognuna di esse può valere 0, 1 o 2 punti. Le domande aperte sono infatti pensate per verificare se il bambino attribuisca spontaneamente credenze o desideri per spiegare le azioni svolte dal protagonista della storia (e.g.: “Simone cerca i pattini nel baule perché *pensa* che siano lì”: 2 punti), se si limiti a una spiegazione più legata agli aspetti situazionali (e.g.: “Simone cerca i pattini nel baule perché *li ha messi* lì”: 1 punto) oppure se dia risposte incoerenti o errate (zero punti);
- punteggio totale (max 113 punti): dato dalla somma del quantitativo e del qualitativo.

La presenza del punteggio qualitativo è un aspetto peculiare del *ToM Storybooks*; esso permette un’interessante approfondimento della competenza di ToM che il bambino possiede, perché la risposta a una domanda “sì/no” può essere data correttamente ma a seguito di un ragionamento sbagliato.

La versione olandese del test permette anche il calcolo di un punteggio standardizzato, il ToMQ, con media 100 e d.s. 15, ottenuto stimando la curva dei punteggi attraverso un metodo di *smoothing* non-lineare sui dati grezzi (Blijd-Hoogewys et al., 2008).

Concretamente, il test si presenta come un serie di sei libri illustrati, che consentono al bambino di giocare aprendo finestrelle, accarezzando peluche, applicando faccine magnetiche con le espressioni emotive del protagonista. Il test risulta molto piacevole e

questo è un aspetto importante, poiché permette anche ai bambini più piccoli di seguire agevolmente la somministrazione individuale, che è abbastanza lunga: essa richiede infatti 40-50 minuti, con una pausa obbligatoria dopo i primi 3 libri.

Riassumendo, il *ToM Storybooks* presenta caratteristiche di particolare interesse:

- valuta l'insieme delle competenze implicate nella comprensione della mente, propria e altrui, secondo l'imposta teorica di Wellman (1990);
- è uno strumento particolarmente innovativo dal punto di vista della costruzione, poiché permette di ottenere misure diverse di una medesima competenza, riflettendo in questo modo una valutazione più stabile e affidabile della competenza sottostante (Hughes e Dunn, 1998; Hughes et al., 2000);
- è concepito per analizzare lo sviluppo e la mutua connessione fra i differenti aspetti della comprensione della mente in un periodo di tempo piuttosto ampio (3-8 anni), durante il quale queste competenze compaiono e maggiormente si sviluppano;
- poiché è stato pensato per la somministrazione a bambini anche molto piccoli, molta cura è stata data all'accessibilità e all'attrattiva del test;
- ha dimostrato buona affidabilità: l'alpha di Cronbach sugli item dicotomici è .90; il K di Cohen sulla codifica alle risposte aperte è alta ( $K=.97-.99$ ); l'affidabilità test-retest è buona sia sui bambini a sviluppo tipico ( $r=.86$ ,  $p<.001$ ) che sui bambini con PDD-NOS ( $r=.98$ ; Blijd-Hoogewys, et al., 2008);
- ha dimostrando una buona validità discriminante rispetto a bambini con disturbo pervasivo dello sviluppo (PDD-NOS) ad alto funzionamento (Blijd-Hoogewys, et al., 2008).

## 2.2. La procedura

Abbiamo tradotto in italiano la versione “Sam” del *ToM Storybooks*, a partire dalla versione inglese messa a punto dagli autori stessi; l'adeguatezza della traduzione italiana

dello strumento è stata verificata tramite una *back-translation* in olandese del testo italiano (van de Vijver e Hambleton, 1996), successivamente discussa con i colleghi olandesi.

Per le modalità di somministrazione, si è seguita la procedura indicata dagli autori olandesi. La raccolta dei dati è avvenuta tra maggio 2005 e novembre 2010. I partecipanti sono stati reclutati tramite asili nido, scuole d'infanzia ed elementari di Torino e provincia e, in alcuni casi, tramite contatti personali dei ricercatori. Nel caso di collaborazione con i servizi educativi, gli insegnanti consegnavano alle famiglie della loro classe il foglio di consenso informato alla partecipazione alla ricerca e, in allegato, il foglio per la raccolta dei dati anagrafici, in forma anonima. I bambini sono stati osservati individualmente presso la scuola frequentata, in un'aula e in un momento della giornata deciso in accordo con gli insegnanti, dopo aver ottenuto l'autorizzazione dei genitori. La somministrazione del test è sempre stata preceduta da un periodo di familiarizzazione tra gli osservatori e i bambini. Nel caso dei contatti personali, il rapporto con la famiglia e la consegna del consenso informato e dei dati anagrafici sono stati curati direttamente dai ricercatori e i bambini hanno svolto il test a casa loro, in una stanza tranquilla, oppure presso il Laboratorio di Osservazione del Comportamento del Dipartimento di Psicologia di Torino. Cinque bambini (4 maschi e 1 femmina) di età media di 39 mesi, sono stati esclusi dall'analisi dei dati, perché non hanno portato a termine il test.

### 2.3. Il campione

I dati sono relativi a un campione di 446 bambini, bilanciato per genere (223 femmine), di età compresa fra 2 anni e mezzo e 8 anni e mezzo (per l'età media e la numerosità ad ogni fascia, si veda Tabella 2). Nella fascia di età dei 4 anni, vi è la maggiore numerosità di bambini, proprio nel periodo in cui è previsto teoricamente il maggiore cambiamento qualitativo nella comprensione della ToM.

Tabella 2

Punteggi medi in relazione all'età (campione italiano, N = 446)

N	Età media		Punteggi ToM Storybooks		
	In anni	In mesi (DS)	Quantitativo (DS)	Qualitativo (DS)	Totale (DS)
52	3 anni	39 (2)	34.19 (7.00)	.42 (.78)	34.62 (7.44)
116	4 anni	48 (3)	39.89 (8.24)	3.31 (4.04)	43.20 (11.41)
70	5 anni	61 (3)	49.27 (10.27)	8.41 (5.91)	57.69 (15.42)
90	6 anni	72 (3)	54.56 (7.64)	11.29 (5.56)	65.84 (12.45)
65	7 anni	85 (3)	59.28 (6.36)	14.98 (5.34)	74.26 (10.74)
53	8 anni	96 (3)	63.30 (6.56)	17.62 (5.30)	80.92 (11.16)

Dalla descrizione del livello socio-economico sono stati escluse informazioni di 17 bambini, 37 madri e 50 padri, poiché i genitori non hanno riportato in modo completo i dati anagrafici. Duecentottanta bambini (62.7%) hanno fratelli, e 90 di essi sono primogeniti. Le madri hanno un'età media di circa 38 anni (min=24, max=49, DS=5) e i padri di circa 41 anni (min=26, max=64, DS=6). Lo 81.6% delle madri e lo 81.8% dei padri sono nati in Italia (in particolare, il 60.5% delle madri e il 56.3% dei padri sono nati in Torino o provincia, il 7.2% delle madri e il 6.7% dei padri nel Nord Italia). Lo 8.3% delle madri e il 5.1% dei padri sono nati in Europa, mentre il 5.4 delle madri e il 6.1% dei padri sono nati al di fuori della Comunità Europea. Sebbene tutti i bambini del campione padroneggiassero adeguatamente la lingua italiana, 76 di loro (16.1% del campione) hanno almeno un genitore straniero e, per 30 di essi (6.8%), i genitori riportano che, in casa, si parla una lingua straniera in modo esclusivo o in alternativa all'italiano.

Il livello di formazione scolastica dei genitori è abbastanza elevato: sono circa 12 gli anni medi di studio di entrambi i genitori ( $DS=4$  anni), corrispondenti alla conclusione del ciclo di media superiore. In specifico, il 36.5% delle madri e il 32.7% dei padri hanno un diploma di media superiore; il 34.5% delle madri e il 38.5% dei padri hanno la licenza media inferiore; il 19.3% delle madri e il 15.9% dei padri sono laureati; e solo il 1.3% delle madri e il 1.6% dei padri posseggono la licenza elementare.

I genitori svolgono prevalentemente un'attività impiegatizia (36.7% delle madri e 28% dei padri), operaia (23.5% delle madri e 40% dei padri) o commerciale (6.5% delle madri e 9.2% dei padri); il 12.6% delle madri sono casalinghe, mentre la condizione di disoccupazione riguarda marginalmente questo campione (4.7% delle madri e 1.6% dei padri).

#### **2.4. Analisi dei dati**

Per valutare l'affidabilità del ToM Storybooks, abbiamo calcolato l'Alfa di Cronbach sui 77 item dicotomici per ogni fascia di età, per tenere sotto controllo possibili differenze dovute all'età dei bambini. Sulle codifiche delle risposte alle domande aperte è stato calcolato il K di Cohen per l'accordo tra giudici.

Per controllare la possibile influenza delle variabili demografiche sui punteggi quantitativo, qualitativo e totale, abbiamo utilizzato l'Anova (variabili: genere; presenza di fratelli categorizzata in sì/no; provenienza dei genitori categorizzata in "italiana" vs. "almeno un genitore straniero"; lingua parlata in casa categorizzata in "solo italiano", vs "italiano e altro"). Quando sono emerse differenze di età significative tra i gruppi, abbiamo tenuto sotto controllo l'influenza dell'età, attraverso una regressione gerarchica, utilizzando l'età in mesi al primo blocco, e l'altra variabile indipendente presa in considerazione al secondo blocco; per snellire la lettura dei risultati, sono presentate solo le analisi di regressione sul punteggio totale, poiché quelle sui punteggi quantitativo e qualitativo sono

risultate assolutamente analoghe. Poiché in letteratura le differenze tra fratelli vengono rinvenute principalmente rispetto alla Falsa Credenza, abbiamo verificato se fossero presenti differenze in base alla presenza di fratelli anche sul punteggio parziale delle prove di FC presenti nel ToM Storybooks (FC Esplicita e FC Classica; punteggio: min=0, max=9).

L'andamento dei punteggi per età è stato controllato tramite l'Anova, con il test post-hoc di Duncan per controllare che i punteggi dei bambini appartenenti a fasce di età diverse si differenziassero significativamente. Sono stati calcolati i punteggi centili per età.

### 3. Risultati

#### 3.1 Affidabilità dei punteggi

La consistenza interna del ToM Storybooks, calcolata sui 77 item dicotomici, è buona: l'alfa di Cronbach varia tra .69 e .88 e ha una media di .80, paragonabile a quella della versione olandese del test.

La codifica delle 18 risposte aperte è stata effettuata da due valutatori (tra cui uno dei due autori). Abbiamo svolto il controllo interosservatore calcolando il K di Cohen sul punteggio attribuito alle 18 risposte aperte, punteggio che può variare da 0 a 2 per risposta. Il controllo è stato svolto sulle risposte di 40 soggetti estratti a caso; i valori dei K sono tutti significativi ( $p < .001$ ), presentano una media di .89 e variano da un massimo di 1.00 a un minimo di .534. I disaccordi sono stati discussi e risolti. La codifica degli unici due item che presentavano un K inferiore a .85 è stata rivista su tutto il campione.

#### 3.2 Controlli sull'influenza delle variabili demografiche

Il genere non influisce sulla prestazione dei bambini al *ToM Storybooks* (Anova univariata, GdL=445; punteggio *quantitativo*: maschi=50.18, femmine=48.35,  $F=2.44$ ,  $p=.119$ ; *qualitativo*: maschi=9.26, femmine=8.31,  $F=1.84$ ,  $p=.176$ ; *totale*: maschi=59.45, femmine=56.65,  $F=2.34$ ,  $p=.127$ ). Poiché le bambine sono significativamente più piccole dei bambini (Età media in mesi: Bambine:  $m=63$ ,  $DS=19$ , Bambini:  $m=67$ ,  $DS=19$ ; t-test:  $t=2.05$ ,  $p<.05$ ), abbiamo verificato che non vi fossero differenze di genere anche



controllando l'età (Regressione gerarchica): l'incremento del modello che comprende il genere è irrilevante (si veda Tabella 3).

Tabella 3

Regressione gerarchica sui punteggi del ToM Storybooks

Variabile indipendente	Variabile dipendente	Modello	Cambiamento di $R^2$	Sig.	Beta	t	Sig.
Genere (N=446)	Punteggio totale	Età in mesi	.000	.879	.793	27.27	<.001
		Genere			.004	.15	.879
Presenza di fratelli (N=429)	Punteggio totale	Età in mesi	.001	.287	.789	26.56	<.001
		Fratelli			.032	1.07	.287
	Punteggio FC	Età in mesi	.000	.983	.490	11.50	<.001
		Fratelli			-.001	-.021	.983

I bambini che hanno fratelli presentano una prestazione migliore al test (Anova univariata, GdL=428; punteggio *quantitativo*: senza fratelli=46.95, con fratelli=50.61,  $F=8.52$ ,  $p<.005$ ; *qualitativo*: senza fratelli=7.50, con fratelli=9.52,  $F=7.19$ ,  $p<.01$ ; *totale*: senza fratelli=54.44, con fratelli=60.13,  $F=8.51$ ,  $p<.005$ ). Tuttavia, controllando l'età, questa differenza scompare (regressione gerarchica: l'incremento del modello che comprende i fratelli è irrilevante, si veda Tabella 3). Infatti, i bambini con fratelli sono significativamente più vecchi dei figli unici (67 mesi rispetto a 62 mesi; t-test:  $t=-2.86$ ,  $p<.005$ ) ed è questo che determina il loro punteggio maggiore. Rispetto al punteggio parziale di FC non emergono differenze significative (Anova, GdL=428; senza fratelli=3.83, con fratelli=4.18,  $F=1.88$ ,  $p>.10$ ), anche controllando per l'età (regressione gerarchica): l'incremento del modello che comprende i fratelli è irrilevante (si veda Tabella 3).

Non sono emerse differenze significative sui punteggi dei bambini in base al luogo di provenienza dei genitori (Anova univariata, GdL=409; punteggio *quantitativo*: genitori italiani=49.28, un genitore straniero=49.94,  $F=.16$ ,  $p>.20$ ; *qualitativo*: genitori italiani=8.69,

un genitore straniero=9.34;  $F=.43$ ,  $p>.20$ ; *totale*: genitori italiani=57.96, un genitore straniero=59.28,  $F=.26$ ,  $p>.20$ ).

Anche prendendo in considerazione la lingua che viene parlata nella famiglia dei bambini, non sono emerse differenze tra i punteggi (Anova univariata,  $GdL=412$ ; punteggio *quantitativo*: italiano=49.19, italiano e altro=50.03,  $F=.13$ ,  $p>.20$ ; *qualitativo*: italiano=8.67, italiano e altro=10.47,  $F=1.60$ ,  $p>.20$ ; *totale*: italiano=57.86, italiano e altro=60.50,  $F=.52$ ,  $p>.20$ ).

### 3.3 Andamento dei punteggi e punteggi standardizzati

I punteggi complessivi della standardizzazione italiana mostrano un chiaro incremento con l'età (si veda Tabella 2). L'Anova univariata sul punteggio totale, con il confronto multiplo post-hoc di Duncan (sottoinsieme per  $\alpha=.05$ ), evidenzia la presenza di 6 sottoinsiemi: ogni età si caratterizza quindi come avente un punteggio che significativamente si differenzia da quello delle altre età. Si noti anche come la prestazione dei bambini presenti un incremento maggiore tra i 4 e i 5 anni (si veda la Figura 1). I punteggi centili sono presentati in Tabella 4.

Figura 1

Andamento dei punteggi medi per età (campione italiano, N=446)

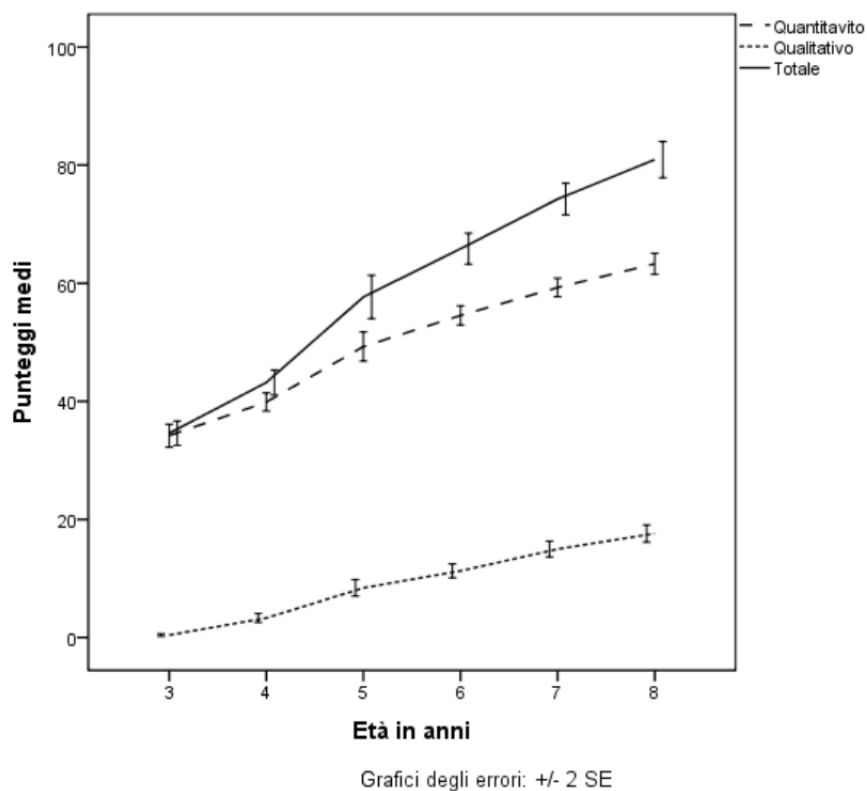


Tabella 4

Punteggi in centili alle diverse età (N = 446)

	Percentili						
	5	10	25	50	75	90	95
3 anni	23	27	30	34	40	46	49
4 anni	27	32	35	42	48	60	66
5 anni	31	37	48	59	67	81	86
6 anni	42	51	59	65	75	85	87
7 anni	55	58	67	73	84	87	89
8 anni	56	65	77	83	88	94	96

#### 4. Discussione

Il ToM Storybooks è un test innovativo nel panorama degli strumenti di valutazione della ToM, per diverse ragioni:

- esso permette di osservare l'evoluzione della ToM a cavallo tra l'età prescolare e scolare, con particolare attenzione alle competenze più precoci di ToM;
- è pensato per valutare questa competenza multidimensionalmente, toccando i diversi aspetti che la compongono (dal riconoscimento delle emozioni, alla distinzione mentale-fisico, alla psicologia del desiderio-credenza, ecc.);
- prevede misure composite, poiché la maggior parte delle prove viene proposta più volte, in contesti di storie differenti, permettendo una più precisa stima della competenza del bambino;
- e, non da ultimo, permette di osservare, attraverso le domande aperte, il ragionamento alla base delle risposte del bambino, garantendo quindi una migliore valutazione della competenza.

Questi aspetti di interesse, relativi al modo in cui il test è stato costruito, sono inoltre accompagnati da buone caratteristiche psicometriche. Come risulta dai dati preliminari più sopra presentati, l'affidabilità del ToM Storybooks nella sua versione italiana è buona: l'alpha sui 77 item dicotomici è di .80 e il K di Cohen sui 18 item a risposta aperta è .89, valori paragonabili a quelli ottenuti nella versione olandese del test. Lo strumento si è rivelato, inoltre, molto soddisfacente rispetto all'andamento delle risposte del campione italiano: i punteggi crescono infatti in modo regolare con l'età, e tutte le fasce di età sono ben differenziate. L'incremento maggiore dei punteggi si verifica tra i 4 e i 5 anni, come ci aspettiamo che sia, in accordo con la teoria sullo sviluppo della ToM, che identifica appunto intorno ai 4 anni la comprensione della falsa credenza.

Non sono emerse differenze di genere, così come in altri studi presenti in letteratura. L'apparente differenza osservata tra i punteggi dei bambini con e senza fratelli può essere

spiegata dalla diversa età media dei due gruppi, dato che i bambini con fratelli sono significativamente più grandi. Questo dato è in contrasto con quelli rilevati negli studi più recenti, sebbene in letteratura siano presenti risultati discordanti (Das e Babu, 2004; Wang e Su, 2009). E' però possibile una interpretazione di questa differenza: la presenza di fratelli sembra legata a migliori prestazioni principalmente in relazione al superamento della FC (McAlister e Peterson, 2007), mentre nel *ToM Storybooks* la ToM è valutata in modo più ampio, multidimensionale, prendendo in considerazione le diverse componenti che la costituiscono, e questo aspetto metodologico potrebbe essere alla base di una non concordanza dei risultati.

Il ToM Storybooks, inoltre, non è risultato sensibile ad altri aspetti socio-demografici, come la provenienza straniera dei genitori dei bambini e, soprattutto, il fatto che in famiglia venga parlata una lingua diversa dall'italiano, in modo esclusivo oppure in alternativa ad esso. Quest'ultimo elemento ci sembra particolarmente importante, perché esclude l'interferenza di aspetti relativi alle differenze linguistiche nella risposta alle prove: il ToM Storybooks si presenta così come uno strumento adatto anche alla nuova realtà multietnica e multiculturale della società italiana.

I dati qui presentati sono i risultati preliminari di un ampio lavoro di ricerca, volto alla validazione italiana dello strumento oltre che alla sua standardizzazione. Stiamo attualmente lavorando alla validazione: due progetti sono in corso, uno relativo alla somministrazione a un campione di bambini con Disturbo Pervasivo dello Sviluppo (Bulgarelli, Molina e Salomone, *in stampa*), l'altro relativo alla validità concorrente con il T.E.C.: cf. Molina, Bulgarelli, Arati (2008) per una presentazione dei dati preliminari. E' inoltre in corso l'ampliamento del campione ad età superiori agli 8 anni, principalmente per permettere un confronto con i punteggi dei bambini con sviluppo atipico, che sono presi in carico dai servizi sanitari anche in tarda fanciullezza. Un confronto sull'andamento dei

punteggi del campione italiano e di quello olandese, e la creazione del punteggio ToM-Q per la popolazione italiana secondo la metodologia dello *smoothing*, sono in preparazione.

Infine, stiamo realizzando, in collaborazione con l'Università di Pavia, una versione dello strumento per i bambini non vedenti

## Bibliografia

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "Theory of Mind"? *Cognition*, 21, 37-46.

Battistelli, P. (1997). *Io penso che tu pensi ... Le origini della comprensione della mente*. Milano: Franco Angeli.

Blijd-Hoogewys, E. M. A., van Geert, P. L. C., Serra, M., & Minderaa, R. B. (2008). Measuring Theory of Mind in children. Psychometric Properties of the ToM Storybooks. *Journal of Autism and Developmental Disorders*, 38, 1907-1930.

Bulgarelli D., Molina P., & Salomone E. (*in stampa*). Valutare la competenza emotiva: significato e strumenti nello sviluppo tipico e atipico. In: I. Grazzani & C. Riva Crugnola (Ed.), *Lo sviluppo della competenza emotiva: percorsi tipici e atipici*. Milano: UNICOPLI.

Charman, T., Ruffman, T., & Clements, W. (2002). Is there a gender difference in false belief development?. *Social Development*, 11, 1-10.

Das, S., & Babu, N. (2004). Children's Acquisition of a Theory of Mind: The Role of Presence vs Absence of Sibling. *Psychological Studies*, 49, 36-44.

Fischer, K. W., & Paré-Blagoev, J. (2000). From Individual Differences to Dynamic Pathways of Development. *Child Development*, 71, 850-853.

Flavell, J. H., Miller P. H., & Miller, S. (1993). *Cognitive development*. Englewood Cliffs, NJ: Prentice-Hall.

- Ford, D. H., & Lerner, R. M. (1995). *Teoria dei sistemi evolutivi: un approccio integrato* (S. Castelli, a cura di). Milano: Cortina.
- Happé, F. G. E. (1994). An advanced test of theory of mind: Understanding of story characters' thought and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders*, 24, 129–154.
- Hughes, C., Adlam, A., Happé, F., Jackson, J., Taylor, A., & Caspi A. (2000). Good test-retest reliability of standard and advanced false belief tasks across a wide range of abilities. *Journal of Child Psychology and Psychiatry*, 41, 483-490.
- Hughes, C., & Dunn, J. (1998). Understanding mind and emotion: Longitudinal associations with mental-state talk between young friends. *Developmental Psychology*, 34, 1026-1037.
- Kovacs, A. M. (2009). Early bilingualism enhances mechanisms of false-belief reasoning. *Developmental Science*, 12, 48-54.
- McAlister, A., & Peterson, C. (2007). A longitudinal study of child siblings and theory of mind development. *Cognitive Development*, 22, 258-270.
- Milligan, K., Astington J.W., & Dack, L. A. (2007). Language and Theory of Mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development*, 78, 622-646.
- Mitchell, P. (1997). *Introduction to theory of mind. Children, autism and apes*. London: Arnold.
- Molina, P., Bulgarelli, D. (2008). La distinzione fra fisico e mentale in età prescolare e scolare: alcuni dati dal test Tom Storybooks. In: C. Fiorilli & O. Albanese (Eds.), *I processi di conoscenza dei bambini: credere, pensare, conoscere* (pp. 55-68). Azzano san Paolo (BG): Junior.

- Molina, P., Bulgarelli, D., Arati, L. (2008). Andamento di sviluppo nelle risposte dei bambini al TEC e al ToM Storybooks: uno studio di validità concorrente. In: O. Albanese e P. Molina, (Eds.), *Lo sviluppo della comprensione delle emozioni e la sua valutazione* (pp. 160-166). Milano: UNICOPLI.
- Muris, P., Steerneman, P., Meesters, C., Merckelbach, H., Horselenberg, R., van den Hogen, T., & van Dongen, L. (1999). The ToM Test: A new instrument for assessing Theory of Mind in normal children and children with Pervasive Developmental Disorders. *Journal of Autism and Developmental Disorders*, 29, 67-80.
- Renouf, A., Brendgen, M., Parent, S., Vitaro, F., Zelazo, P. D., Boivin, M., Dionne, G., Tremblay, R. E., Perusse, D., & Seguin, J. R. (2010). Relations between theory of mind and indirect and physical aggression in kindergarten: Evidence of the moderating role of prosocial behaviors. *Social Development*, 19, 535-555.
- Steerneman, P., Meesters, C., & Muris, P. (2002). *Tom-Test*. Leuven-Apeldoorn: Garant.
- Tager-Flusberg, H. (2003). Exploring the relationship between theory of mind and social-communicative functioning in children with autism. In B. Repacholi & V. Slaughter (Eds.), *Individual differences in theory of mind. Implications for typical and atypical development* (pp. 197-212). Hove: Psychology Press.
- Thelen, E., & Smith, L. B. (1994). *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, Mass.: The MIT Press
- Thommen E., Cartier-Nelles B., Wiesendanger S., Guidoux A., & Blijd-Hoogewys, E. M. A. (2010, Ottobre). Theory of mind in children with Autistic Spectrum Disorders: a longitudinal study. Poster presentato al *IX International Congress AUTISM-EUROPE*, Catania, I.
- van de Vijver, F., & Hambleton, R. K. (1996). Translating Tests: Some Practical Guidelines. *European Psychologist*, 1, 89-99.



- van Geert, P. L. C. (2003). Dynamic Systems Approaches and Modeling of Developmental Processes. In: J. Valsiner, & K. J. Conolly (Eds.), *Handbook of Developmental Psychology* (pp. 640-672). London: Sage.
- Vesterinen, J. (2008) *The ToM Storybooks as a tool of studying children's Theory of Mind in Finland* (Tesi di laurea non pubblicata). Università di Jyväskylä, Jyväskylä, SF.
- Walker, S. (2005). Gender Differences in the Relationship Between Young Children's Peer-Related Social Competence and Individual Differences in Theory of Mind. *The Journal of Genetic Psychology Research and Theory on Human Development*, 166, 297-312.
- Wang, Y., & Su, Y. (2009). False belief understanding: Children catch it from classmates of different ages. *International Journal of Behavioral Development*, 33, 331-336.
- Wellman H. M. (1990). *The child's theory of mind*. Cambridge, London: The MIT Press.
- Wellman H. M., Cross D, & Watson J. (2001). Meta-analysis of Theory of Mind development: The truth about False Belief. *Child Development*, 72, 655–684.
- Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, 75, 523–541.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding and deception. *Cognition*, 13, 103-1